Applicants' representatives pointed out that the Minami teaches away from the claimed invention and that it is improper to combine references where the references teach away from their combination.

Rejections under 35 U.S.C. §103

Claims 1-27 have been rejected under 35 U.S.C. § 103 as being obvious in light of a combination of Minami in view of Suzuki et al. (U.S. Patent No. 5,440, 639, hereinafter, Suzuki), Cooper (U.S. Patent No. 4,910,779, hereinafter, Cooper) or Tanaka (U.S. Patent No. 5,498,478, hereinafter, Tanaka).

The Office Action asserts that it would have been obvious to combine the stereo voice transmission apparatus of Minami with the head related features or functions of Suzuki, Cooper or Tanaka. The Office Action states that Minami teaches a two-way communication system or conference system which provides virtual audio or stereo voice transmission for the listener, and asserts that it would have been obvious to replace the transfer function "imparted" in Minami with the head related features or functions of Suzuki, Cooper or Tanaka. To establish a prima facie case of obviousness there must be some suggestion or motivation, either in the references themselves or in the knowledge generally available to one of ordinary skill in the art, to modify the reference or to combine reference teachings. Claim 1 recites a conference station comprising right and left spatially disposed microphones connected to a communication channel for receiving right and left audio signals, **wherein the differences between the right and left audio signals represent a head-related transfer function**". According to claim 1, the differences between the right and left audio signals represent a head-related transfer function. Applicants point out that there is no motivation to modify the Minami reference or to combine the teachings of cited references. In fact, it is improper to combine references where the references teach away from their combination. In re Grasselli, 713 F.2d 731, 743, 218 USPQ 769, 779 (Fed. Cir. 1983). Adding a head-related transfer function to the Minami reference destroys the intended use of the reference *i.e.*, reducing the cost of transmitting a stereo voice signal across a telephone line. Furthermore, Minami prevents

fluctuation of sound image localization by transmitting a monaural voice transmission when one person is speaking and transmitting a stereo voice transmission when more than one person is speaking. This method teaches away from imparting the head-related-transfer function/spatial components to the monaural voice signal.

Minami discloses an apparatus where "stereo voice transmission is performed in the multiple simultaneous utterance mode, and **monaural voice transmission is performed in a single utterance mode.**" As pointed out in the background of Minami (column 2, lines 50-58), when a monaural voice transmission is directly applied to a multiple utterance, the sound image localization has an undesirable fluctuation. Minami's objective is to prevent fluctuations of the sound image localization, while reducing the cost of transmission, by transmitting the single and multiple utterances by different transmission methods and/or different transmission rates.

Minami is transmitting a single utterance as a monaural voice signal and is transmitting multiple simultaneous utterances as a stereo voice transmission to realize a high band compression ratio. During a conversation when one person is speaking (*i.e.*, single utterance), Minami transmits the speech as a monaural voice signal (column 13, lines 39-65). Minami does not alter this signal with any spatial information before transmission. During a conversation when more than one person is speaking (*i.e.*, multiple utterances), Minami transmits the speech as a stereo voice transmission. Minami does not alter this signal with any additional spatial information before transmission.

The Office Action states that the transfer function "imparted" by Minami may not be disclosed as a head-related-transfer function. However, Applicants point out that the transfer function in Minami does not "impart" anything to the transmitted signal, *i.e.*, the transfer function is a ratio of the input signal and the output signal (column 13, lines 35-65). In the Minami reference, a single voice signal is transmitted when one person is speaking, and the transfer function is used once the signal is received to reproduce the originally recorded signals for the listener. The transfer function **does not "impart" any spatial components** on the voice signal. In fact, the Minami reference "teaches away" from adding a head-related transfer function. The intended use of Minami, *i.e.*, reducing the cost of transmitting a stereo voice signal across a telephone line, is contrary to

imparting a head-related transfer function. Adding a head-related transfer function to the Minami reference is counter productive. Imparting a head-related-transfer function to the monaural signal would increase the cost of transmission. Furthermore in Minami, when a multiple utterance is detected and a stereo transmission is utilized, the original signal is not altered by a transfer function to impart spatial components. There are no spatial sound information being applied to the original signal. (column 9, lines 45-59)

Additionally, Minami shows adding sound localization information to a received location. This spatial information is not imparted before transmission. The sound image localization information is obtained from a sound image localization control information generator (column 15, line8 - column 16 line 21). The sound image localization control information generator generates a plurality of pieces of sound image localization control information L_k including, as information, at least one of delay, phase, and gain differences determined in correspondence with the position of the image displayed on the screen. When the plurality of pieces of sound image localization control information L_k are used, for example, as shown in FIG. 11, sound image localization control is performed as if a voice is produced from the position of speaker's mouth of the image 712 on the screen 710. More specifically, the screen 710 is divided into N x M blocks, and sound image localization is controlled in units of blocks. Even when any one of the delay, phase, and gain differences is used, or a combination of the differences is used, the above sound image localization control can be performed.

Therefore, the combination of the cited references does not teach or suggest that **the differences between the right and left audio signals represent a head-related transfer function**. Applicants respectfully request that the examiner reconsider the combination of references asserted against the claim and withdraw the rejection under 35 U.S.C. § 103.

Claims 2-12 depend directly or indirectly from claim 1 and, as such, include the limitations thereof. Applicants submit that the differences between what is set forth in claims 2-12 including the limitations of claim 1 from which they depend, are such that they would not have been obvious

to those of ordinary skill in the art in view of the cited references. Withdrawal of the rejection under 35 U.S.C. § 103 is respectfully requested.

Claim 13 recites converting audio information into right and left audio signals at a conference station, wherein the **conversion imparts** a differential characteristic to the right and left audio signals, and the **differential characteristic is represented by a head-related transfer function**, and the right and left audio signals comprise spatialized audio and **transmitting audio information representative of said spatialized audio from the conference station across a communication channel to a remote station**. As stated above, adding a head-related transfer function to the Minami reference destroys the intended use of the reference *i.e.,* reducing the cost of transmitting a stereo voice signal across a telephone line. Furthermore, Minami prevents fluctuation of sound image localization by transmitting a monaural voice transmission when one person is speaking and transmitting a stereo voice transmission when more than one person is speaking. This method teaches away from imparting the head-related-transfer function/spatial components to the monaural voice signal.

Minami discloses an apparatus where "stereo voice transmission is performed in the multiple simultaneous utterance mode, and **monaural voice transmission is performed in a single utterance mode**." As pointed out in the background of Minami (column 2, lines 50-58), when a monaural voice transmission is directly applied to a multiple utterance, the sound image localization has an undesirable fluctuation. Minami's objective is to prevent fluctuations of the sound image localization, while reducing the cost of transmission, by transmitting the single and multiple utterances by different transmission methods and/or different transmission rates.

Minami is transmitting a single utterance as a monaural voice signal and is transmitting multiple simultaneous utterances as a stereo voice transmission to realize a high band compression ratio. During a conversation when one person is speaking (*i.e.,* single utterance), Minami transmits the speech as a monaural voice signal (column 13, lines 39-65). Minami does not alter this signal with any spatial information before transmission. During a conversation when more than one person

is speaking (*i.e.*, multiple utterances), Minami transmits the speech as a stereo voice transmission. Minami does not alter this signal with any additional spatial information before transmission.

The Office Action states that the transfer function "imparted" by Minami may not be disclosed as a head-related transfer function. However, Applicants point out that the transfer function in Minami does not "impart" anything to the transmitted signal, *i.e.*, the transfer function is a ratio of the input signal and the output signal (column 13, lines 35-65). In the Minami reference, a single voice signal is transmitted when one person is speaking, and the transfer function is used once the signal is received to reproduce the originally recorded signals for the listener. The transfer function **does not "impart" any spatial components** on the voice signal. In fact, the Minami reference "teaches away" from adding a head-related transfer function. The intended use of Minami, *i.e.*, reducing the cost of transmitting a stereo voice signal across a telephone line, is contrary to imparting a head-related transfer function. Adding a head-related transfer function to the Minami reference is counter productive. Imparting a head-related-transfer function to the monaural signal would increase the cost of transmission. Furthermore in Minami, when a multiple utterance is detected and a stereo transmission is utilized, the original signal is not altered by a transfer function to impart spatial components. There are no spatial sound information being applied to the original signal. (column 9, lines 45-59)

Additionally, Minami shows adding sound localization information to a received location. This spatial information is not imparted before transmission. The sound image localization information is obtained from a sound image localization control information generator (column 15, line8 - column 16 line 21). The sound image localization control information generator generates a plurality of pieces of sound image localization control information L_k including, as information, at least one of delay, phase, and gain differences determined in correspondence with the position of the image displayed on the screen. When the plurality of pieces of sound image localization control information L_k are used, for example, as shown in FIG. 11, sound image localization control is performed as if a voice is produced from the position of speaker's mouth of the image 712 on the screen 710. More specifically, the screen 710 is divided into N x M blocks, and sound image

localization is controlled in units of blocks. Even when any one of the delay, phase, and gain differences is used, or a combination of the differences is used, the above sound image localization control can be performed.

Therefore, the combination of the cited references does not teach or suggest that **the differences between the right and left audio signals represent a head-related transfer function**. Applicants respectfully request that the examiner reconsider the combination of references asserted against the claim and withdraw the rejection under 35 U.S.C. § 103.

Claim 14 depends directly from claim 13 and, as such, includes the limitations thereof. Applicants submit that the differences between what is set forth in claim 14 including the limitations of claim 13 from which it depends, is such that it would not have been obvious to those of ordinary skill in the art in view of the cited references. Withdrawal of the rejection under 35 U.S.C. § 103 is respectfully requested.

Claim 15 recites a **head-related transfer function unit connected to the communications system for imparting a head-related transfer function** to the audio signal **to produce a spatialized audio signal**". As stated above, adding a head-related transfer function to the Minami reference destroys the intended use of the reference *i.e.,* reducing the cost of transmitting a stereo voice signal across a telephone line. Furthermore, Minami prevents fluctuation of sound image localization by transmitting a monaural voice transmission when one person is speaking and transmitting a stereo voice transmission when more than one person is speaking. This method teaches away from imparting the head-related-transfer function/spatial components to the monaural voice signal.

Minami discloses an apparatus where "stereo voice transmission is performed in the multiple simultaneous utterance mode, and **monaural voice transmission is performed in a single utterance mode.**" As pointed out in the background of Minami (column 2, lines 50-58), when a monaural voice transmission is directly applied to a multiple utterance, the sound image localization has an undesirable fluctuation. Minami's objective is to prevent fluctuations of the sound image

localization, while reducing the cost of transmission, by transmitting the single and multiple utterances by different transmission methods and/or different transmission rates.

Minami is transmitting a single utterance as a monaural voice signal and is transmitting multiple simultaneous utterances as a stereo voice transmission to realize a high band compression ratio. During a conversation when one person is speaking (*i.e.,* single utterance), Minami transmits the speech as a monaural voice signal (column 13, lines 39-65). Minami does not alter this signal with any spatial information before transmission. During a conversation when more than one person is speaking (*i.e.,* multiple utterances), Minami transmits the speech as a stereo voice transmission. Minami does not alter this signal with any additional spatial information before transmission.

The Office Action states that the transfer function "imparted" by Minami may not be disclosed as a head-related transfer function. However, Applicants point out that the transfer function in Minami does not "impart" anything to the transmitted signal, *i.e.,* the transfer function is a ratio of the input signal and the output signal (column 13, lines 35-65). In the Minami reference, a single voice signal is transmitted when one person is speaking, and the transfer function is used once the signal is received to reproduce the originally recorded signals for the listener. The transfer function **does not "impart" any spatial components** on the voice signal. In fact, the Minami reference "teaches away" from adding a head-related transfer function. The intended use of Minami, *i.e.,* reducing the cost of transmitting a stereo voice signal across a telephone line, is contrary to imparting a head-related transfer function. Adding a head-related transfer function to the Minami reference is counter productive. Imparting a head-related-transfer function to the monaural signal would increase the cost of transmission. Furthermore in Minami, when a multiple utterance is detected and a stereo transmission is utilized, the original signal is not altered by a transfer function to impart spatial components. There are no spatial sound information being applied to the original signal. (column 9, lines 45-59)

Additionally, Minami shows adding sound localization information to a received location. This spatial information is not imparted before transmission. The sound image localization information is obtained from a sound image localization control information generator (column 15,

line8 - column 16 line 21). The sound image localization control information generator generates a plurality of pieces of sound image localization control information L_k including, as information, at least one of delay, phase, and gain differences determined in correspondence with the position of the image displayed on the screen. When the plurality of pieces of sound image localization control information L_k are used, for example, as shown in FIG. 11, sound image localization control is performed as if a voice is produced from the position of speaker's mouth of the image 712 on the screen 710. More specifically, the screen 710 is divided into N x M blocks, and sound image localization is controlled in units of blocks. Even when any one of the delay, phase, and gain differences is used, or a combination of the differences is used, the above sound image localization control can be performed.

Therefore, the combination of the cited references does not teach or suggest that **the differences between the right and left audio signals represent a head-related transfer function**. Applicants respectfully request that the examiner reconsider the combination of references asserted against the claim and withdraw the rejection under 35 U.S.C. § 103.

Claims 16-18 depend directly from claim 15 and, as such, include the limitations thereof. Applicants submit that the differences between what is set forth in claims 16-18 including the limitations of claim 15 from which they depend, are such that they would not have been obvious to those of ordinary skill in the art in view of the cited references. Withdrawal of the rejection under 35 U.S.C. § 103 is respectfully requested.

Claim 19 recites transmitting the audio signal from the **transmitting station to a spatial sound conference bridge and imparting a head-related transfer function to the audio signal to create a spatialized audio signal** and sending the spatialized audio signal from the spatial sound conference bridge to a receiving station. As stated above, adding a head-related transfer function to the Minami reference destroys the intended use of the reference *i.e.,* reducing the cost of transmitting a stereo voice signal across a telephone line. Furthermore, Minami prevents fluctuation of sound image localization by transmitting a monaural voice transmission when one person is speaking and transmitting a stereo voice transmission when more than one person is speaking. This

method teaches away from imparting the head-related-transfer function/spatial components to the monaural voice signal.

Minami discloses an apparatus where "stereo voice transmission is performed in the multiple simultaneous utterance mode, and **monaural voice transmission is performed in a single utterance mode**." As pointed out in the background of Minami (column 2, lines 50-58), when a monaural voice transmission is directly applied to a multiple utterance, the sound image localization has an undesirable fluctuation. Minami's objective is to prevent fluctuations of the sound image localization, while reducing the cost of transmission, by transmitting the single and multiple utterances by different transmission methods and/or different transmission rates.

Minami is transmitting a single utterance as a monaural voice signal and is transmitting multiple simultaneous utterances as a stereo voice transmission to realize a high band compression ratio. During a conversation when one person is speaking (*i.e.,* single utterance), Minami transmits the speech as a monaural voice signal (column 13, lines 39-65). Minami does not alter this signal with any spatial information before transmission. During a conversation when more than one person is speaking (*i.e.,* multiple utterances), Minami transmits the speech as a stereo voice transmission. Minami does not alter this signal with any additional spatial information before transmission.

The Office Action states that the transfer function "imparted" by Minami may not be disclosed as a head-related transfer function. However, Applicants point out that the transfer function in Minami does not "impart" anything to the transmitted signal, *i.e.,* the transfer function is a ratio of the input signal and the output signal (column 13, lines 35-65). In the Minami reference, a single voice signal is transmitted when one person is speaking, and the transfer function is used once the signal is received to reproduce the originally recorded signals for the listener. The transfer function **does not "impart" any spatial components** on the voice signal. In fact, the Minami reference "teaches away" from adding a head-related transfer function. The intended use of Minami, *i.e.,* reducing the cost of transmitting a stereo voice signal across a telephone line, is contrary to imparting a head-related transfer function. Adding a head-related transfer function to the Minami reference is counter productive. Imparting a head-related-transfer function to the monaural signal

would increase the cost of transmission. Furthermore in Minami, when a multiple utterance is detected and a stereo transmission is utilized, the original signal is not altered by a transfer function to impart spatial components. There are no spatial sound information being applied to the original signal. (column 9, lines 45-59)

Additionally, Minami shows adding sound localization information to a received location. This spatial information is not imparted before transmission. The sound image localization information is obtained from a sound image localization control information generator (column 15, line8 - column 16 line 21). The sound image localization control information generator generates a plurality of pieces of sound image localization control information L_k including, as information, at least one of delay, phase, and gain differences determined in correspondence with the position of the image displayed on the screen. When the plurality of pieces of sound image localization control information L_k are used, for example, as shown in FIG. 11, sound image localization control is performed as if a voice is produced from the position of speaker's mouth of the image 712 on the screen 710. More specifically, the screen 710 is divided into N x M blocks, and sound image localization is controlled in units of blocks. Even when any one of the delay, phase, and gain differences is used, or a combination of the differences is used, the above sound image localization control can be performed.

Therefore, the combination of the cited references does not teach or suggest that **the differences between the right and left audio signals represent a head-related transfer function**. Applicants respectfully request that the examiner reconsider the combination of references asserted against the claim and withdraw the rejection under 35 U.S.C. § 103.

Claims 20 and 21 depend directly from claim 19 and, as such, include the limitations thereof. Applicants submit that the differences between what is set forth in claims 20 and 21 including the limitations of claim 19 from which they depend, are such that they would not have been obvious to those of ordinary skill in the art in view of the cited references. Withdrawal of the rejection under 35 U.S.C. § 103 is respectfully requested.

Claim 22 recites a method for **conducting a spatial sound conference** comprising the steps of receiving an audio signal at a transmitting station, transmitting the audio signal from the transmitting station to a receiving station, **imparting a head-related transfer function to the audio signal to create spatialized audio signal**, and playing the spatialized audio signal on spatially disposed loudspeakers in the receiving station. As stated above, adding a head-related transfer function to the Minami reference destroys the intended use of the reference *i.e.,* reducing the cost of transmitting a stereo voice signal across a telephone line. Furthermore, Minami prevents fluctuation of sound image localization by transmitting a monaural voice transmission when one person is speaking and transmitting a stereo voice transmission when more than one person is speaking. This method teaches away from imparting the head-related-transfer function/spatial components to the monaural voice signal.

Minami discloses an apparatus where "stereo voice transmission is performed in the multiple simultaneous utterance mode, and **monaural voice transmission is performed in a single utterance mode**." As pointed out in the background of Minami (column 2, lines 50-58), when a monaural voice transmission is directly applied to a multiple utterance, the sound image localization has an undesirable fluctuation. Minami's objective is to prevent fluctuations of the sound image localization, while reducing the cost of transmission, by transmitting the single and multiple utterances by different transmission methods and/or different transmission rates.

Minami is transmitting a single utterance as a monaural voice signal and is transmitting multiple simultaneous utterances as a stereo voice transmission to realize a high band compression ratio. During a conversation when one person is speaking (*i.e.,* single utterance), Minami transmits the speech as a monaural voice signal (column 13, lines 39-65). Minami does not alter this signal with any spatial information before transmission. During a conversation when more than one person is speaking (*i.e.,* multiple utterances), Minami transmits the speech as a stereo voice transmission. Minami does not alter this signal with any additional spatial information before transmission.

The Office Action states that the transfer function "imparted" by Minami may not be disclosed as a head-related transfer function. However, Applicants point out that the transfer

function in Minami does not "impart" anything to the transmitted signal, *i.e.*, the transfer function is a ratio of the input signal and the output signal (column 13, lines 35-65). In the Minami reference, a single voice signal is transmitted when one person is speaking, and the transfer function is used once the signal is received to reproduce the originally recorded signals for the listener. The transfer function **does not "impart" any spatial components** on the voice signal. In fact, the Minami reference "teaches away" from adding a head-related transfer function. The intended use of Minami, *i.e.*, reducing the cost of transmitting a stereo voice signal across a telephone line, is contrary to imparting a head-related transfer function. Adding a head-related transfer function to the Minami reference is counter productive. Imparting a head-related-transfer function to the monaural signal would increase the cost of transmission. Furthermore in Minami, when a multiple utterance is detected and a stereo transmission is utilized, the original signal is not altered by a transfer function to impart spatial components. There are no spatial sound information being applied to the original signal. (column 9, lines 45-59)

Additionally, Minami shows adding sound localization information to a received location. This spatial information is not imparted before transmission. The sound image localization information is obtained from a sound image localization control information generator (column 15, line8 - column 16 line 21). The sound image localization control information generator generates a plurality of pieces of sound image localization control information L_k including, as information, at least one of delay, phase, and gain differences determined in correspondence with the position of the image displayed on the screen. When the plurality of pieces of sound image localization control information L_k are used, for example, as shown in FIG. 11, sound image localization control is performed as if a voice is produced from the position of speaker's mouth of the image 712 on the screen 710. More specifically, the screen 710 is divided into N x M blocks, and sound image localization is controlled in units of blocks. Even when any one of the delay, phase, and gain differences is used, or a combination of the differences is used, the above sound image localization control can be performed.

Therefore, the combination of the cited references does not teach or suggest that **the differences between the right and left audio signals represent a head-related transfer function**. Applicants respectfully request that the examiner reconsider the combination of references asserted against the claim and withdraw the rejection under 35 U.S.C. § 103.

Claim 23 depends directly from claim 22 and, as such, includes the limitations thereof. Applicants submit that the differences between what is set forth in claim 23 including the limitations of claim 22 from which it depends, is such that it would not have been obvious to those of ordinary skill in the art in view of the cited references. Withdrawal of the rejection under 35 U.S.C. § 103 is respectfully requested.

Claim 24 recites **a spatial sound conference bridge** comprising at least two input ports for receiving at least two audio signals and at least two audio signal output ports, a head-related transfer function unit connected to at least of said input ports for **imparting a head-related transfer function** to a corresponding audio signal **to produce at least one spatialized audio signal**, and at least two output ports connected to the head-related transfer function unit for transmitting the spatialized audio signal. As stated above, the transfer function of Minami **does not "impart" any spatial components** on the voice signal. In fact, adding a head-related transfer function to the Minami reference is contrary to Minami's intended use, *i.e.*, reducing the cost of transmitting a stereo voice signal across a telephone line by reducing the bandwidth. Adding a head-related transfer function to the Minami reference is counter productive. Therefore, the combination of the cited references does not teach or suggest **a spatial sound conference bridge that imparts a head-related transfer function** to at least one received audio signal **to produce at least one spatialized audio signal**. Applicants respectfully request that the examiner reconsider the combination of references asserted against the claim and withdraw the rejection under 35 U.S.C. § 103.

Claims 25 and 26 depend directly or indirectly from claim 24 and, as such, include the limitations thereof. Applicants submit that the differences between what is set forth in claims 25 and 26 including the limitations of claim 24 from which they depend, are such that they would not have

been obvious to those of ordinary skill in the art in view of the cited references. Withdrawal of the rejection under 35 U.S.C. § 103 is respectfully requested.

Claim 27 recites a method for conducting **a spatial sound conference** comprising the steps of receiving at least two monaural audio signals, **generating at least two sets of spatialized audio signals** from the at least two monaural audio signals **using at least two head-related transfer functions**, compiling at least one composite signal from the at least two sets of spatialized audio signals, transmitting at least one composite signal to a location, and playing at least one composite signal at the location. As stated above, the transfer function of Minami **does not "impart" any spatial components** on the voice signal. In fact, adding a head-related transfer function to the Minami reference is contrary to Minami's intended use, *i.e.*, reducing the cost of transmitting a stereo voice signal across a telephone line by reducing the bandwidth. Adding a head-related transfer function to the Minami reference is counter productive. Therefore, the combination of the cited references does not teach or suggest **generating at least two sets of spatialized audio signals** from the at least two monaural audio signals **using at least two head-related transfer functions**. Applicants respectfully request that the examiner reconsider the combination of references asserted against the claim and withdraw the rejection under 35 U.S.C. § 103.

CONCLUSION

This application is in condition for allowance, and early notice of same is earnestly solicited. Should the examiner have any questions, comments or suggestions in furtherance of the prosecution of this application, he is invited to contact applicants' representative by telephone at the number indicated below.

Respectfully submitted,

Date: 22 Sept 99

Mark Ungerman
Registration No. 32,070

Fulbright & Jaworski L.L.P.
801 Pennsylvania Avenue, N. W.
Washington, D. C. 20004-2615
Telephone: (202)662-0200

565301.1